1974

Accuracy of the chi-square approximation for 2 x 3 contingency tables with very small expected frequencies /

Mei Tung P. Huang
Lehigh University

ACCURACY OF THE CHI-SQUARE APPROXIMATION FOR 2 X 3 CONTINGENCY TABLES WITH VERY SMALL EXPECTED FREQUENCIES

by

Mei Tung P. Huang

A Thesis
Presented to the Graduate Committee
of Lehigh University
in Candidacy for the Degree of
Master of Science
in the
School of Education
Division of Educational Measurement and Research

Lehigh University

1974

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

May 1, 1974
(date)

David L. Marsh
Professor in Charge

Paul Van R. Miller
Department Chairman

ACKNOWLEDGMENT

The author wishes to express her gratitude to Dr. Miller, Chairman of the Department of Educational Measurement and Research and Dr. March, Professor of Educational Measurement and Research, for their introduction of this subject and their kind assistance and encouragement; also the computing center for using their valuable time.

Sincere thanks are also due to Mrs. D. Fielding who typed this manuscript.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The chi-square test of independence for contingency tables is often used to test for a relation between variables in the population being studied. There is a universal consensus that the chi-square test is a good approximation to the exact test when the average expectations are greater than 5. The purpose of this study was to investigate the accuracy of the chi-square approximations compared to the exact cumulative probabilities in 2 X 3 contingency tables which had one or more expected frequencies less than 1.

Under the limitation that one or more expected frequencies should be smaller than 1, 936 different sets of marginal totals were generated for sample sizes from 6 through 18. In addition, 522 and 673 sets of marginal totals for sample sizes 24 and 30 were also included. The exact and chi-square probabilities were computed for all 2 X 3 contingency tables which were generated from those sets of marginal totals. The chi-square probabilities were compared with the corresponding exact cumulative probabilities by three different ways. First, the range and the mean absolute percentage errors of the chi-square probabilities were computed. Second, the number of times that the chi-square probabilities underestimated or overestimated the exact probabilities were calculated. Third, the extent of agreement between chi-square and exact probabilities with respect to accepting or

rejecting a null hypothesis at .01 and .05 levels of significance was investigated.

The error range of chi-square probability was very large. The smaller the sample is, the narrower the range. The mean absolute percentage errors range from 38 to 68. The higher means of the absolute percentage errors occurred at both small and large sample sizes. The relative accurate chi-square approximations are obtained when the average expected frequencies are between 2 and 3. However, if close approximation to the exact probabilities are needed, the chi-square test is still poor.

The disagreements between chi-square and exact probabilities in most cases were due to the underestimates of chi-square to exact probabilities. The ratios of underestimations to overestimations increases along with the increase of sample size. Therefore, the accuracy of chi-square approximation increases as the sample increases.

With respect to the extent of agreement, in over 90 percent of the tables studied the chi-square probability lead to the same conclusion as the exact probability in accepting or rejecting a null hypothesis at both the .01 and .05 levels.

CHAPTER I

CHI-SQUARE AND EXACT TESTS FOR CONTINGENCY TABLES

Introduction

Contingency Table: When the observations on two qualitative variables or one qualitative and one quantitative variable are classified into a two-way table, they are known as contingency data and the table as a contingency table. The classifications are made according to two major but independent characteristics with several levels. As the sample is randomly drawn from an infinite and homogeneous population, each member in the sample is distributed into one of the cells of the contingency table by its relation to one of the several levels within both characteristics.

Chi-square Test

The chi-square test is one of the non-parametric tests which is used as a test of significance when we have data that are in terms of percentages or proportions and that can be reduced to frequencies. Any of the applications of chi-square test have to do with discrete data. However, any continuous data may be reduced to categories and the data so tabulated that chi-square may be applied.

<u>Usage of Chi-square Test.</u> The major usage of the chi-square test for contingency tables as indicated by Downie and Heath (1970) is to test the hypothesis concerning the significance of the differences of of the responses of two or more groups to a stimulus of one type or another. In this case, the chi-square test is referred to as a test of independence of the variables in the population being studied.

<u>Chi-quare Test of R X C Contingency Table</u>. In order to test the hypothesis that there is no relation between the variables in the population, the chi-square test is used to compare the observed results with the expected frequencies.

The null hypothesis can be tested by first calculating the statistic:

$$X^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \qquad (1)$$

Here, O_{ij} is the observed number of cases categorized in the ith row of jth column. E_{ij} is the number of cases expected under the hypothesis to be categorized in the ith row of jth column. $\sum_{i=1}^{r} \sum_{j=1}^{c}$ is the sum of all R rows and C columns, i.e., the sum of all cells.

If all of the expected frequencies are large, the distribution of X^2 is reasonably close to the chi-square distribution with (r - 1) (c - 1) degrees of freedom. For this reason, the X^2 test is often referred to as a chi-square test.

Therefore, the mathematical definition of the chi-square test is the sum of the ratios of the square of the differences between the observed and expected frequencies to expected frequencies in each cell. However, in order to find the expected frequencies of each cell, the products of the two marginal totals common to a particular cell is computed. Then, the result is divided by the total number of cases. The formula is:

$$M_{ij} = r_i \, c_j \qquad E_{ij} = \frac{M_{ij}}{N} \tag{2}$$

If the observed frequencies are in close agreement with the expected frequencies, the differences will of course be small. Consequently, the value of X^2 is small. With a small value of X^2, we may not reject the null hypothesis that there is no difference between the observed frequencies and the expected frequencies. However, if some or many of the differences are large, the value of X^2 will also be large. The larger the value of chi-square, the more likely it is that the two variables are not independent.

Exact Test for R X C Contingency Table

For computing the exact test of significance for a contingency table, Fisher developed a special formula for computing the exact probability for a 2 x 2 contingency table. Later, Freeman and Haltman (1951) derived a general method for computing the exact probability for a contingency table with K different and independent classifications. The derivation of their formula was based on four assumptions:

1. The population is homogeneous and infinite.
2. The sample is done with the replacement of the sampled items.
3. The sampling is random.
4. The marginal totals are considered fixed in repeated sampling.

If all of the above assumptions are satisfied, the exact probability (P_X) of obtaining the observed array of cell frequencies X (X_{ij}) in a sample size N with two different and independent classifications: A and B with R and C classes respectively is:

$$P_X = \frac{\prod^{r} (r_i) \prod^{c} (c_j)}{N! \prod_{i=1}^{r} \prod_{j=1}^{c} (X_{ij})} \qquad (3)$$

Minimum Expectations for the Chi-square Test

Common Agreements and Theories of Small Expectations. The chi-square test of independence for a contingency table is simpler than the exact test. There is common agreement that the chi-square test is a good approximation to the exact test when the sample is large. This implies that the chi-square test may yield poor approximation to the exact test when the expected frequencies are small. As to what the definition of small should be, no consensus has been reached. However, one widely used rule of thumb is that each cell should have an expectation not less than 5.

According to Cochran (1954), the chi-square test can be used only if fewer than 20 percent of the cells have expected frequencies of less than 5 and if no cell has an expected frequencies of less than 1 in a contingency table.

Tate and Clelland (1959) mentioned that if the degrees of freedom is more than 2 and the average expectation is more than 5, then each cell does not necessarily have to have an expectation more than 5.

Walker and Lev (1953) suggested that if roughly approximate probabilities are acceptable, an expectation of only two in a cell is sufficient. However, if the expectations in all of the cells but one are 5 or more, an expectation of only 1 in the remaining cell is sufficient to provide a fair approximation to the exact probabilities.

In studying the robustness of the chi-square test of independence for contingency tables in skew and uniform populations, Roscoe and Byars (1970) found that the chi-square test of independence is very robust in both situations. As to the minimum expectations for use of the chi-square test, they concluded that the chi-square test gives an excellent approximation with average expected frequencies as low as 2 if the data is drawn fron an uniform population. However, if the data is drawn from a skew population, there is a tendency for chi-square test to become conservative with small samples.

According to the common agreement, if none of the expected frequencies are less than 5, X^2 will follow the chi-square distribution with (r - 1)(k - 1) degrees of freedom. Otherwise, the chi-square probability will be a poor approximation to the corresponding cumulative exact probability. However, this is only an assumption.

Accuracy of Chi-square Approximation with Small Expectations. In order to investigate the accuracy of the chi-square probabilities when compared to the exact probabilities for 2 x 3 contingency table with average expectations smaller than 5, March (1970) compared 2,845 different sets of marginal totals which were generated from values of N from 6 to 30. In his study, one restriction was imposed. Tables which had one or more expected frequencies less than one were eliminated.

His findings show a steady improvement of the chi-square approximation as the sample size, and thus the average expected frequency increases. However, if close approximations to the exact probabilities are desired, the chi-square test might at times be poor with small cell expectations.

His second finding is that if one is only interested in accepting or rejecting some hypotheses at a special level of significance, the X^2 test leads to the same decision as the exact test at the .05 level of significance better than 95 percent of the time if the average expectation is 3 or larger and none of the expectations are less than 1.

In the area of overestimates or underestimates compared to the exact test, the chi-quare test is biased toward overestimation in the .005 - .010 probability region and toward underestimation in the .010 - .200 region. The relative amount of underestimation decreases as the sample size increases.

March's study suggests the tendency for the ratios of underestimates to overestimates of chi-square to increase as the exact probability increases and decrease as the sample size increases. Moreover, the tentative conclusion about the patterns of underestimation and overestimation for a limited number of exact distribution suggests that a worthwhile correction might exist for the chi-square test when applied to contingency tables. Therefore, further research is needed on the numbers and patterns of underestimates and overestimates for contingency tables with expectations less than 1.

Purpose of this Study

There is a consensus in the literature that the chi-square test gives a poor approximation to the exact probability when one or more expected frequencies is less than 1. In order to do further research in this area, this study investigated the accuracy of the chi-square approximation in 2 x 3 contingency tables which had one or more expected frequencies less than 1. Tables that had expected frequencies larger than 1 in all of the cells were eliminated from this study.

CHAPTER II

PROCEDURE FOR COMPUTATION

Introduction

This study is concerned with the accuracy of chi-square approximations to exact probabilities when the expected frequencies are less than 1 in one or more cells in 2 x 3 contingency tables. Therefore, one limitation was imposed in the procedure of computation. Tables which had expectations more than 1 in all of the cells were eliminated.

There are three main stages in the computational procedures. They were the generation of marginal totals for forming appropriate tables, the computation of exact and chi-square probabilities for all 2 x 3 contingency tables that could be generated under the restriction imposed, and the comparisons of the results based on the absolute percentage differences and on the extent of agreement at both the .01 and .05 probability levels.

The Selection of Sample Size

The first step was to decide the minimum and maximum size of the sample in order to obtain meaningful results. When the sample size is less than 6, the value of at least one marginal total is equal to zero. Thus, a meaningful selection of sample size has to start from 6. The maximum sample size for this study was 18, which is, of course, an arbitrary number.

In order to see the common trend as the sample size was increasing, this study also investigated the results rendered by sample size 24 and 30. Therefore, comparisons between chi-square and exact probabilities were made for every sample size from 6 to 18 and additional study included the values of N of 24 and 30.

Generations of Row and Column Marginal Totals

After the sample size had been decided, the next step was the generation of different sets of combinations for row and column marginal totals for each sample size being studied. Two restrictions were imposed here. The first one was the elimination of any set of combinations containing zero as a row or a column marginal total. The second limitation was that the program stopped once the permutation of marginal totals had been detected. It was unnecessary to permute the marginal totals since they yielded the same results in the generation of contingency tables as the original ones.

Generation of Row Marginal Totals. The generation of row marginal totals for each sample size was a matter of addition and subtraction. Letting N be the sample size, the first step was to choose the value of the first number of the initial set as N - 1 and the value of the second number as 1.

Successive sets of row combinations were obtained by subtracting 1 from the first number of the previous set to obtain the first number of the next set and adding 1 to the second number of the previous set to obtain the second number of this next set. This procedure continued as long as the value of the first number was larger than or equal to the value of the second number.

Under the two restrictions previously mentioned, this study generated 73 sets of row marginal totals for sample sizes from 6 through 18. In addition, 12 sets of row marginal totals for sample size 24 and 15 sets of row marginal totals for sample size 30 were generated.

Generation of Column Marginal Totals. The steps used to generate column marginal totals were similar to the procedures for generating row marginal totals. Letting N be the sample size, the value of the first number for the first set began with N - 2. Under the two restrictions imposed here, the values of the other two numbers had to be 1.

Additional sets were obtained by subtracting 1 from the first value to obtain a new first value, say N - M. Then, the second and third values were set to be M - 1 and 1. These values were varied in the same manner as that used for generation of row totals. This process continued until the subtraction of 1 from the first value caused it to be less than N/3.

For this study, 170 sets of column marginal totals for sample sizes from 6 through 18 were generated. Besides, 48 sets for sample size 24 and 75 sets for sample size 30 were also formed.

Formation of Sets of Marginal Totals. After the different sets of row and column marginal totals for each sample size had been generated, each set of row marginal totals was combined with every set of column marginal totals in order to form different sets of row and column marginal totals for each value of N.

Since this study is concentrated on examining the nature of chi-square probability as compared to exact probability when one or more of the expected frequencies is less than 1, one limitation was imposed in the procedure for combining column and row totals. While each set of totals was formed, the expected frequency of each cell was calculated by using formula 2 (page 5). Those sets generating expected cell frequencies larger than 1 in all of their cells were deleted from this study immediately. Only those sets of marginal totals generating tables with expectations less than 1 in at least one of the cells were preserved.

A total of 936 sets of marginal totals were formed for sample sizes from 6 to 18. In addition, 522 sets for sample size 24 and 673 sets for sample size 30 were also included for larger sample study.

Calculation of Chi-square and Exact Probabilities

As each set of marginal totals was formed, an existing subroutine (March, 1970, 1972) immediately generated all possible cell combinations and the corresponding exact and chi-square probabilities for every set of marginal total.

The value of chi-square was obtained by using formula 1, (page 4). Then the computation of the chi-square probability corresponding to the value of chi-square at 2 degrees of freedom was computed by this formula:

$$P\left(X^2\right) = \int_0^{X^2} \frac{1}{2} e^{-\frac{1}{2} y} \, dy = e^{-\frac{1}{2} X^2} \qquad (4)$$

The exact probability was computed by using formula 3 (page 6) derived by Freemand and Halton. The values of the exact probabilities were sorted from minimum to maximum for every set of marginal totals. Then the cumulative exact probabilities were obtained by addition.

Comparisons and Summarizations of the Results

At this stage, the exact cumulative probabilities and the corresponding chi-square probabilities were compared. In this study, 25,654 probability pairs were compared.

There were three types of comparisons. First, the range of the absolute errors and the mean absolute percentage error for each sample size were computed by the following formula:

$$\text{Error} = \frac{P\ (X^2) - P\ (E)}{P\ (E)} \times 100\% \qquad (6)$$

Here $P\ (X^2)$ is the chi-square probability and $P\ (E)$ represents the corresponding exact cumulative probabilities. Second, the number of times the chi-square probabilities overestimated or under-estimated the exact probabilities were calculated for each value of N in order to examine the accuracy of chi-square approximation. Third, a count was made of the number of times that the chi-square and exact probabilities would have led to the same or different decisions at the .01 and .05 levels of significance respectively. The precentages of agreements were also calculated.

Validity of the Program

All of the procedures mentioned above were executed on the CDC 6400 digital computer at Lehigh University.

In order to check the validity of the programs, the generation or row and column marginal totals were verified by hand calculation for some randomly chosen sample sizes. Sets of marginal totals were checked randomly to assure that they did not violate the restriction that the expectation of at least one cell in the contingency table is less than 1.

The best way to examine the accuracy of the exact probabilities was to check the cumulative exact probabilities of obtaining all possible outcomes for a given set of marginal total should be 1.0. It was found that the computation of exact probabilities were accurate to at least six significant digit numbers.

References to the computer programs can be obtained in the Library of the Lheigh University Computing Center.

CHAPTER III

RESULTS AND DISCUSSIONS

Introduction

Although this study excluded all sets of the marginal totals that generated 2 x 3 contingency tables with all expected frequencies greater than or equal to 1 in all of their cells, there were still 936 different sets of marginal totals generated for values of N from 6 through 18. In addition, 522 and 673 sets of marginal totals were generated for sample sizes of 24 and 30. Exact and chi-square probabilities were computed for all 2 x 3 contingency tables which were generated from those sets of marginal totals. After all the probabilities were calculated, the chi-square probabilities were compared with the exact cumulative probabilities.

Types of Comparisons

There were three types of comparisons for every sample size. First, the range and the mean absolute percentage errors of the chi-square probabilities were computed so as to examine the accuracy of chi-square approximations to exact probabilities. Second, the ratios of overestimates and underestimates of chi-square probabilities were also calculated. Third, the extent of agreement between the chi-square and exact probabilities with respect to acceptance or rejection of a null hypothesis at the .01 and .05 level of significance respectively was computed.

Range and Mean Absolute Percentage Errors. Using formula 5 (page 15), the errors of the chi-square approximations for 2 x 3 contingency tables were obtained. Then, the mean percentage error for every sample size was computed. As shown in Table I, the error range for each sample size is very large. The smaller the sample is, the narrower the range. The widest range occurs when the sample size is 30.

When one or more expected frequencies are less than 1, the range of the means of absolute errors is from 38 to 68. However, when all of the expected frequencies are larger than 1 but their average is smaller than 5, the means of absolute percentage errors range from 34 to 99 (March, 1970). The comparisons of the mean percentage errors between the expected frequencies larger than 1 and smaller than 1 are illustrated in Table II and Figure 1.

TABLE I

Means and range of absolute percentage errors for the chi-square approximations to the exact cumulative probabilities for 2 x 3 contingency tables.

N	N of Cases	Means of Absolute Percentage Errors	Range of Absolute Percentage Errors
6	41	56.6%	0% - 85.1%
7	57	49.1	5.7 - 89.4
8	93	45.4	1.1 - 92.7
9	146	43.8	1.4 - 95.0
10	214	40.3	.9 - 96.6
11	303	40.2	.7 - 97.8
12	397	39.5	.1 - 174.6
13	568	38.0	.7 - 178.2
14	704	38.3	.0 - 280.0
15	883	39.6	.0 - 291.8
16	1134	38.4	.1 - 429.1
17	1427	40.3	.1 - 454.4
18	1395	41.7	.0 - 639.1
24	4380	51.6	.0 - 1999.1
30	9532	68.1	.0 - 5973.2

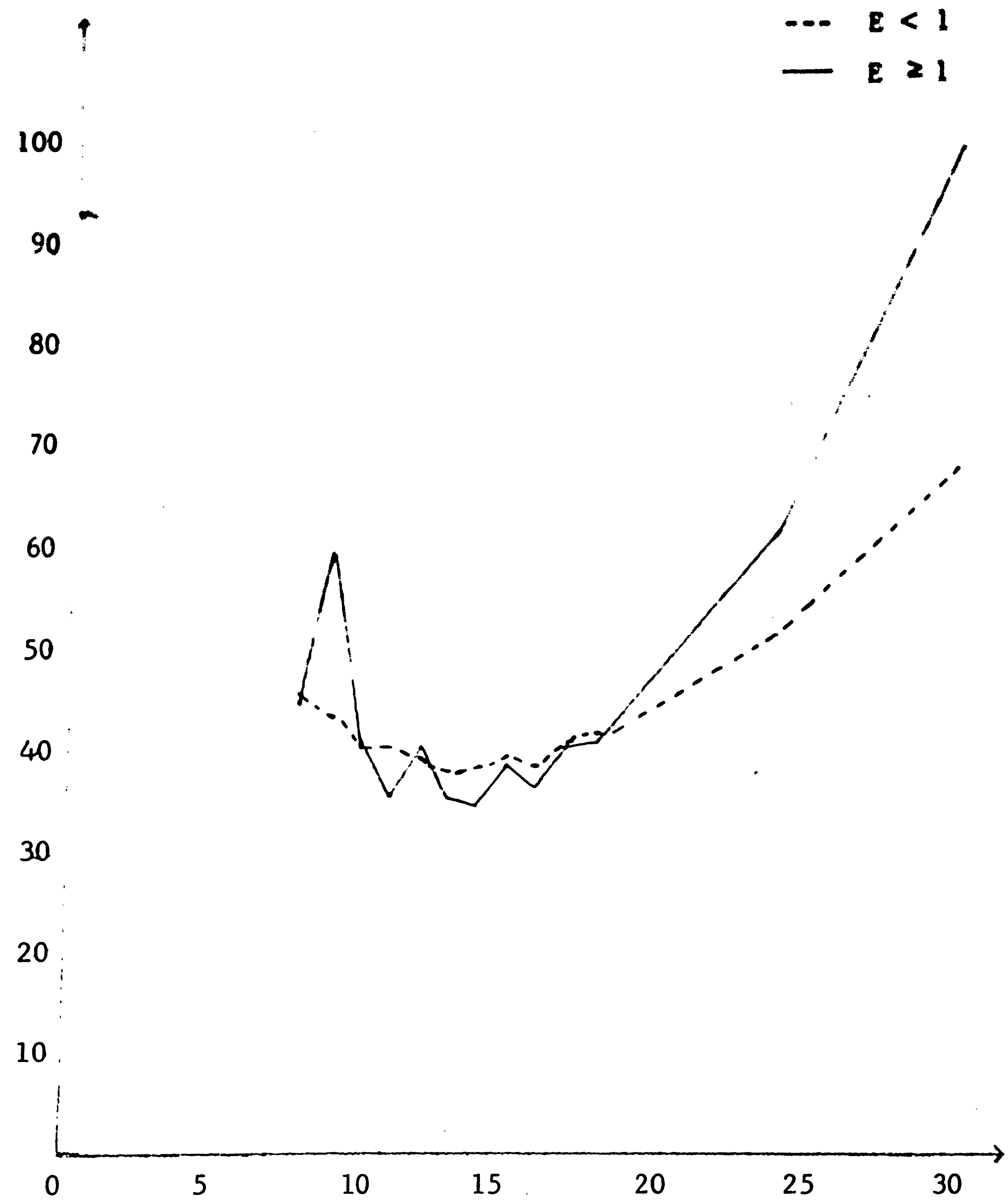

Fig. 1 Comparisons of the means of the absolute percentage errors between all expectations larger than 1 but smaller than 5 and one or more expectations smaller than 1.

TABLE II

Comparisons of the means of absolute percentage errors between all expectations larger than 1 but smaller than 5 and one or more expectations smaller than 1.

N	Expectations Larger than 1 (%)	Expectations Smaller than 1 (%)
8	44.9	45.4
9	59.1	43.8
10	41.0	40.3
11	35.9	40.2
12	40.3	39.5
13	35.5	38.0
14	34.8	38.3
15	38.3	39.6
16	36.5	38.4
17	40.2	40.3
18	40.8	41.7
24	61.2	51.6
30	99.8	68.1

Careful inspection indicates that the higher mean errors are obtained both with small and large samples. The relatively accurate chi-square approximations occur when the average expected frequencies are between 2 and 3 in both cases. The accuracy of chi-square approximation decreases as the sample size increases in both situations.

Comparatively speaking, the accuracy of chi-square approximations to exact probabilities is less accurate when all the expected frequencies are larger than 1 than when one or more expectations are smaller than 1.

Patterns of Underestimates and Overestimates. Each chi-square probability was compared with the corresponding exact probability for every sample size. The number of times that the chi-square probability underestimated or overestimated the exact probabilities was recorded and the ratios of underestimation to overestimation were computed. These are reported in Table III.

In most of the cases, the chi-square probabilities underestimate the exact probabilities. The ratios of underestimates to overestimates decrease as the sample size increases. Therefore, the accuracy of chi-square approximation increases as the sample becomes larger.

TABLE III

Number of times and percentage of overestimates and underestimates for the chi-square approximations to the exact cumulative probabilities for 2 x 3 contingency tables.

N	Percentage of Underestimates (%)	Percentage of Overestimates (%)	Ratios of Underestimates to Overestimates
6	97.6	2.4	40 (40/ 1)
7	94.7	5.3	18 (54/ 3)
8	94.6	5.4	17.6 (88/ 5)
9	93.8	6.2	15.2 (137/ 9)
10	93.0	7.0	13.3 (199/ 15)
11	89.4	10.6	8.5 (271/ 32)
12	87.4	12.6	6.9 (347/ 50)
13	86.3	13.7	6.3 (490/ 78)
14	85.9	14.1	6.1 (605/ 99)
15	81.5	18.5	4.4 (720/ 163)
16	79.6	20.4	3.9 (903/ 231)
17	78.2	21.8	3.6 (1116/ 311)
18	72.8	27.2	2.7 (1015/ 380)
24	68.8	31.2	2.2 (3008/ 1372)
30	64.1	35.9	1.8 (6113/ 3426)

Note:

Number in parenthesis as (underestimates/overestimates) where overestimates include cases of equality.

Extent of Agreement Between Chi-square and Exact Probabilities. The most common usage of the exact test is the determination of the acceptance or rejection of a certain null hypothesis at a specified level of significance. If the chi-square probability leads to the same decision as the exact probability, the chi-square probability can be used in this situation regardless of the size of the percentage errors.

This study examined the extent of agreement between chi-square and exact probabilities at both .01 and .05 levels of significance. Table IV and V report a very satisfactory extent of agreement at both significance levels.

When the sample size is 6 or 7 at the .05 level of significance, the extent of agreement is from 75 to 85 percent. As the sample size becomes larger, the extent of agreement enlarges to over 90 percent. Furthermore, the trend of increase continues with the increase of the sample size.

The extent of agreement is somewhat different at the .01 level. When the sample size is small, like 6, 7 or 8, decisions based on chi-square probabilities agree with those based on exact cumulative probabilities perfectly. However, the percentage of agreement at the .01 level is over 90 percent in all cases which is still relatively higher than the percentage of agreement at the .05 level. Figure 2 shows the comparisons of the extent of agreement between the .01 and .05 levels.

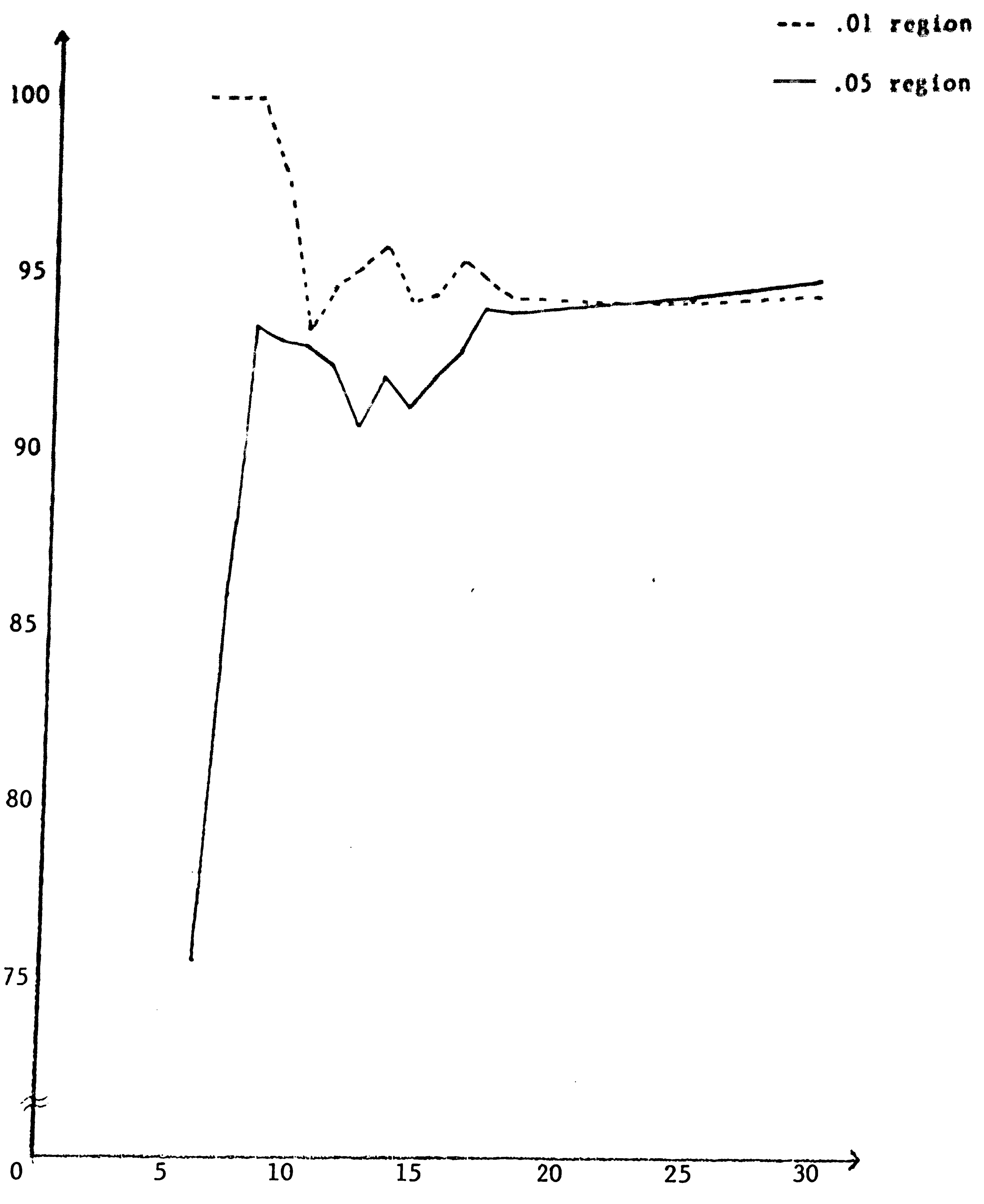

Fig. 2 Comparison of the extent of agreement between chi-square and exact probability for testing significance at the .01 and .05 levels.

The disagreements between chi-square and exact probabilities in most of the cases are due to the tendency for chi-square probabilities to underestimate the exact probabilities. As the sample size increases, the ratios of underestimation to overestimation decreases. Consequently, the extent of agreement increases.

TABLE IV

Extent of agreement between chi-square and exact probabilities at the .05 level of significance.

N	Number of Cases	Number of Agreement	Percentage of Agreement (%)	Number of Disagreement	Percentage of Disagreement (%)
6	41	31	75.6	10	24.4
7	57	49	86.0	8	14.0
8	93	87	93.5	6	6.5
9	146	136	93.2	10	6.8
10	214	199	93.0	15	7.0
11	303	280	92.4	23	7.6
12	397	360	90.7	37	9.3
13	568	524	92.3	44	7.7
14	704	643	91.3	61	8.7
15	883	814	92.2	69	7.8
16	1134	1035	92.9	81	7.1
17	1427	1343	94.1	84	5.9
18	1395	1311	94.0	84	6.0
24	4380	4133	94.4	237	5.6
30	9532	9049	94.9	483	5.1

TABLE V

Extent of agreement between chi-square and exact probabilities at the .01 level of significance.

N	Number of Cases	Number of Agreement	Percentage of Agreement (%)	Number of Disagreement	Percentage of Disagreement (%)
6	41	41	100.0	0	0.0
7	57	57	100.0	0	0.0
8	93	93	100.0	0	0.0
9	146	143	97.9	3	2.1
10	214	200	93.5	14	7.5
11	303	287	94.7	16	5.3
12	397	380	95.2	17	4.3
13	568	544	95.8	24	5.2
14	704	664	94.3	40	5.7
15	883	834	94.5	49	5.6
16	1134	1082	95.4	52	4.6
17	1427	1354	94.9	73	5.1
18	1395	1317	94.4	78	5.6
24	4380	4132	94.3	248	5.9
30	9532	8992	94.3	540	5.8

Table VI and Figure 3 illustrate the comparisons of the extent of agreement at the .05 level between tables which have all of their expected frequencies larger than 1 and tables that have at least one expected frequency smaller than 1. The highest and lowest extent of agreements both occur when all the expectations are larger than 1. The second column in this table shows an increase of the extent of agreement along with the increase of the sample size when all the expectations are larger than 1. In the case where one or more expectations are less than 1, the average extent of agreement is higher but the value of it increases slowly as the sample size increases.

TABLE VI

Comparisons of the extent of agreement at the .05 level between tables which have all expectations larger than 1 and tables that have one or more expected frequencies smaller than 1.

N	Expectations larger than 1 (%)	Expectations smaller than 1 (%)
8	89.5	93.5
9	72.7	93.3
10	86.9	93.0
11	87.9	92.4
12	87.9	90.7
13	91.2	92.3
14	92.2	91.3
15	92.1	92.2
16	92.8	92.9
17	94.8	94.1
18	95.1	94.0
24	96.3	94.4
30	97.3	95.0

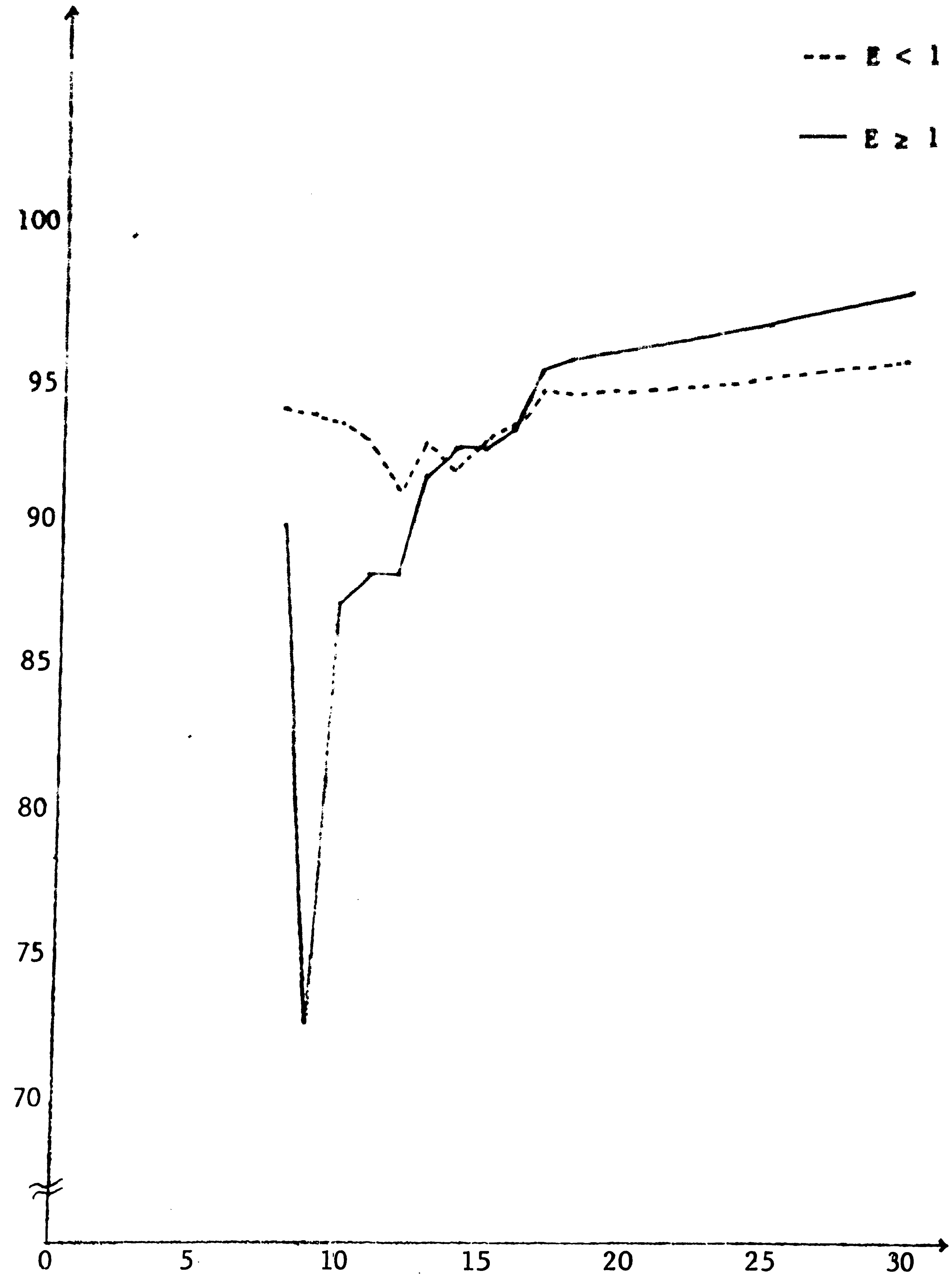

Fig. 3 Comparisons of the extent of agreement at the .05 level between tables which have all expectations larger than 1 and tables that have one or more expected frequencies smaller than 1.

CHAPTER IV

SUMMARY AND CONCLUSIONS

Purpose of Study

The purpose of this study was to investigate the accuracy of chi-square probabilities as compared to the exact cumulative probability in 2 x 3 contingency tables with at least one expected cell frequency smaller than 1.

Computational Procedures

The sample size in this study started at 6, which is the minimum meaningful number for 2 x 3 contingency tables, and ended with 18 which is an arbitrary number. Moreover, the values of N of 24 and 30 were also included in order to see the further trend as the sample size increased.

There were three stages involved in the computations which were performed on CDC 6400 digital computer at Lehigh University.

The first stage was the generation of marginal totals so as to form appropriate tables. One restriction was imposed here. Those marginal totals which generated tables with expected frequencies larger than 1 in all of their cells were eliminated. Under this restriction, 936 different sets of marginal totals were generated for sample sizes from 6 through 18. Besides, 522 sets for N of 24 and 673 sets for N of 30 were generated.

The second stage of computation was the calculation of the chi-square and exact cumulative probabilities for all possible arrays of observed cell frequencies that were generated from a given set of marginal totals. In all 251,654 probability pairs were obtained.

The last stage was the comparisons of the results. There were three kinds of comparisons between chi-square and exact cumulative probabilities. First, the mean absolute percentage errors of chi-square probabilities were computed. Second, the number of times that chi-square underestimated or overestimated the exact probabilities were counted. Third, the extent of agreement between the exact and chi-square probabilities was investigated at both .01 and .05 probability levels. All of these comparisons were made for each sample size.

Conclusion

Recommendations about the minimum expectations to be used in the chi-square test are varied. One widely used rule of thumb indicates that the chi-square test can be used only if the cell expectations are more than 5.

In the study of the accuracy of the chi-square approximation for 2 x 3 contingency tables with average expectations smaller than 5 but all expectations larger than 1, March found the chi-square test can not give accurate approximations to the exact test. However, with respect to accepting or rejecting a null hypothesis, the chi-square probabilities lead to the same decisions as exact probabilities approximately 90 percent of the time.

In this further study of the accuracy of the chi-square approximation for 2 x 3 contingency tables with one or more expected frequencies smaller than 1, both the accuracy of chi-square test to exact probabilities are better than when all expectations are larger than 1.

The range of the absolute percentage errors of the chi-square probabilities is very large. The widest one is from 0 to 5,973 when the sample size is 30. The smaller the sample is, the narrower the error range is. The mean absolute percentage errors, which ranges from 38 to 68, are somewhat different. The higher means of the absolute percentage errors occurred at both small and large sample sizes. The relatively accurate chi-square approximations were obtained when the average expected frequencies are between 2 and 3. As the sample became larger the chi-square test yielded poorer approximations. However, if close approximations to the exact probabilities are needed, the chi-square test is still poor.

As to the extent of agreement betwen chi-square probabilities and exact cumulative probabilities, the finding is very satisfactory. If we are interested in only accepting or rejecting the null hypothesis at the .01 or .05 level of significance, the chi-square probability leads to the same conclusion over 90 percent as often as the exact probability in both levels. The disagreement occurs in most cases due to the underestimations of chi-square to exact probabilities. The ratios of underestimates to overestimates increases as the value of N increases as a whole.

The popular rule of thumb indicates the chi-square test can be used only when all cell expectations are more than 5. In addition, there is consensus that the chi-square test should not be used if any of the expectations are less than 1. According to the findings of this study, it is true that the accuracy of chi-square approximation is poor if close approximation to the exact test is needed when one or more cell expectations are less than 1. This study also showed that the accuracy of chi-square approximation improves as the sample size, and therefore, the average expectation, increases. On the other hand, the empirical results of this study showed over 90 percent agreement between the chi-square test and the exact test with respect to testing a hypothesis at the .01 or .05 levels. Since these results are similar to the findings of March's study (1970) with expected cell frequencies larger or equal to 1 but smaller than 5 in average 2 x 3 contingency tables, the common agreement that the chi-square test should not be used if any of cell expectations are smaller than 1 is no more justified than any other rule of thumb.

However, this study dealt only with 2 x 3 contingency tables having one or more expectations less than 1 in the matrix. Further study is suggested on the effect of having one, two, etc. cell expectations less than 1 in 2 x 3 contingency tables. Caution must be taken in applying the findings of this study to contingency tables with different number of rows and columns. Therefore, further study is also suggested on the accuracy of the chi-square test with small expectations in different kinds of contingency tables.

BIBLIOGRAPHY

Cochran, W. C., "Some Methods for Strengthening the Common X^2 Tests", *Biometrics*, 10:417-451, 1954.

Downie, N. M. and Heath, R. W., "*Basic Statictical Methods*", pp. 196-214, Harper and Row, Publishers, New York, 1970

Freeman, G. H. and Halton, J. H., "Note on an Exact Treatment of Contingency, Goodness of Fit and Other Problems of Significance", *Biometrics*, 38:141-149, 1951.

Guildord, J. P., "*Fundamental Statistics in Psychology and Education*", McGraw Hill Inc., New York, 1965.

Kreft, C. H. and Eeden, C. V., "*A Nonparametric Introduction to Statistics*", The Macmillan Company, New York, 1968.

March, D. L., "Accuracy of the Chi-square Approximation for 2 x 3 Contingency Tables with Small Expectation", An unpublished D. ED., Diss., School of Education, Lehigh University, Bethlehem, Pa., 1970.

March, D. L., "Exact Probabilities for R X C Contingency Tables (G^2)", Communications of the ACM, 15: 991-992, 1972.

Roscoe, J. T. and Byars, J. A., "Some Notes on the Chi-approximation of the Multinominal and Selected Alternatives", Paper Association, Minneapolis, March, 1970.

Tate, M. W. and Clelland, R. C., "*Nonparametric and Shortcut Statistics*", Danville, Illinois, Interstate Printers and Publishers, 1957.

Tate, M. W., "*Statistics in Education and Psychology*", Toronto, Macmillan Company, 1971.

Veldman, D. J., "*Fortran Programming for the Behavioral Sciences*", Holt, Rinehard and Winston, Inc., 1967.

Walker, H. M. and Lev, J., "Statistical Inference", Holt, Rinehard and Winston, Inc., New York, pp. 104-108, 1953.

VITA

Mei Tung P. Huang was born on July 11, 1949 in Hong Kong, the daughter of Mr. and Mrs. Gean Dei Huang. She completed her elementary and secondary education in Taipei, Taiwan.

In June 1971 she received her B. A. in Foreign Language and Literature in National Taiwan University. In September 1972, Miss Huang entered Lehigh University as a major in Educational Measurement and Research.